

DEEP LEARNING COM TENSORFLOW



MODELOS GENERATIVOS

DIEGO RODRIGUES DSC

INENET

CRONOGRAMA

Dia	Aula	Trab
02/09	Workshop de Deep Learning	
04/09	Deep FeedForward	
09/09	Rede Neural Convolutiva	Modelo Baseline
11/09	AutoEncoders	
16/09	Representation & Transfer Learning	Modelo Profundo
18/09	Sequências	
23/09	Modelos Generativos	 Deployment
25/09	Apresentação dos Trabalhos Parte II	

MODELOS GENERATIVOS

- PARTE 1 : TEORIA
 - BUSINESS UNDERSTANDING
 - MODELOS GENERATIVOS EM DEEP LEARNING
 - MODELING
 - MODELOS BAYESIANOS E MARKOVIANOS
 - MÁQUINA DE BOLTZMANN
 - DEEP BELIEVE NETWORK
 - AUTOENCODER
 - GENERATIVE ADVERSARIAL NETWORK
 - TRANSFORMERS
 - DIFUSÃO

PARTE 1 : TEORIA

The image features a light gray background with a subtle gradient. In the top-left and bottom-right corners, there are clusters of realistic water droplets of various sizes, some overlapping. A faint, circular, embossed-like pattern is visible in the upper center of the page.

BUSINESS UNDERSTANDING

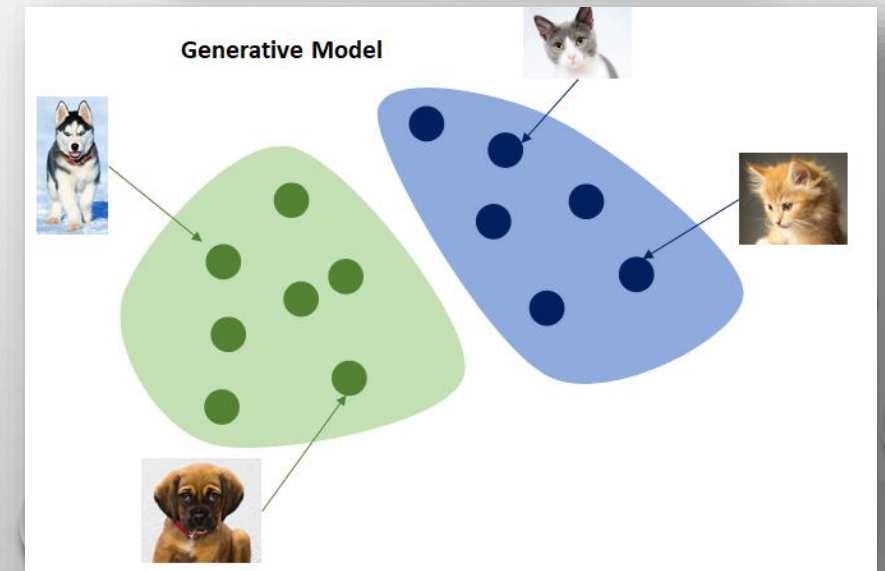
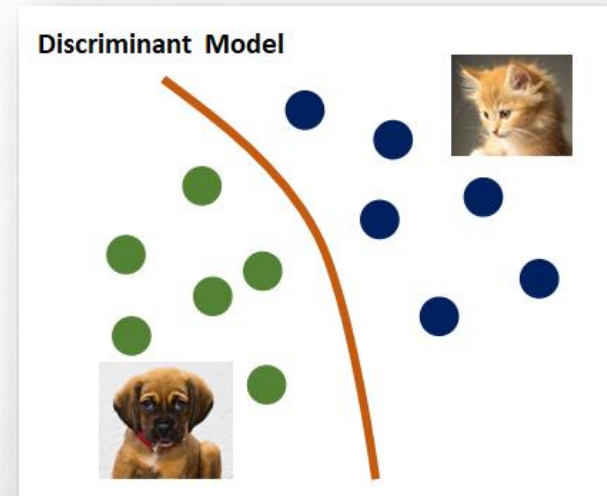
MODELOS GENERATIVOS EM DEEP LEARNING

Modelos generativos são um tipo de modelo de aprendizado de máquina que **aprendem a capturar a distribuição probabilística dos dados de entrada**. O objetivo é **gerar novos dados** que sejam **similares aos dados reais**, amostrando dessa distribuição.

Diferente de **modelos discriminativos** (que classificam), os modelos generativos se concentram em **entender a distribuição dos dados**.

Aplicações

- Criação de conteúdo sintético (imagens, vídeos, áudio).
- Geração de novos exemplos para dados escassos.
- Simulações e melhorias em RL (Reinforcement Learning).



PRINCIPAIS MODELOS GENERATIVOS

1. Autoencoders
2. Variational Autoencoders (VAE)
3. **Generative Adversarial Networks (GAN)**
4. **Modelos de Difusão**

Exemplos

- **DALL-E / MidJourney:** Gera imagens realistas a partir de descrições textuais.
- **DeepMind's WaveNet:** Um modelo gerativo de áudio usado na síntese de voz.
- **GPT:** Modelos como GPT-3 geram texto e completam frases com base em uma distribuição de linguagem.



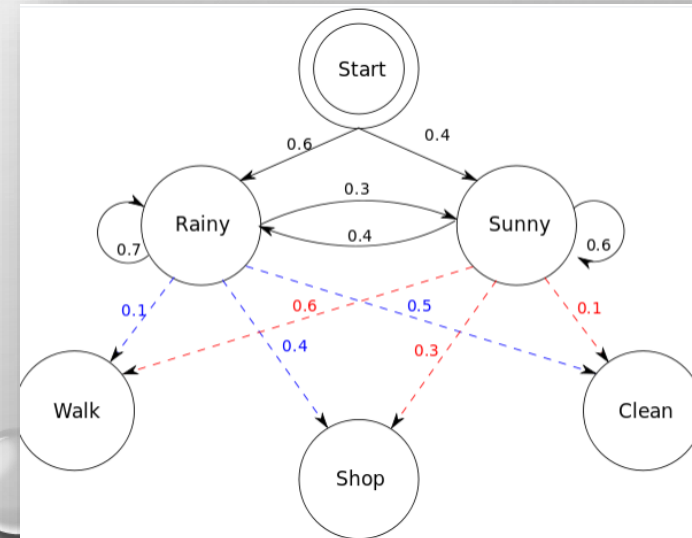
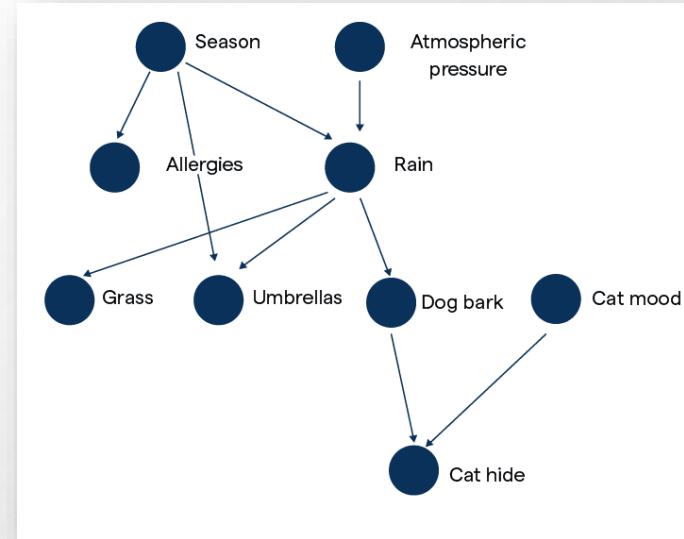
ChatGPT

MODELING

MODELOS MARKOVIANOS E REDES BAYESIANAS 90'S

Durante os anos 90, o **aprendizado de máquina** focava em **técnicas probabilísticas tradicionais**.

Duas das técnicas principais utilizadas para modelar uma Distribuição de Probabilidade de uma Amostra desconhecida eram as **redes bayesianas** e os **modelos de markov ocultos**.

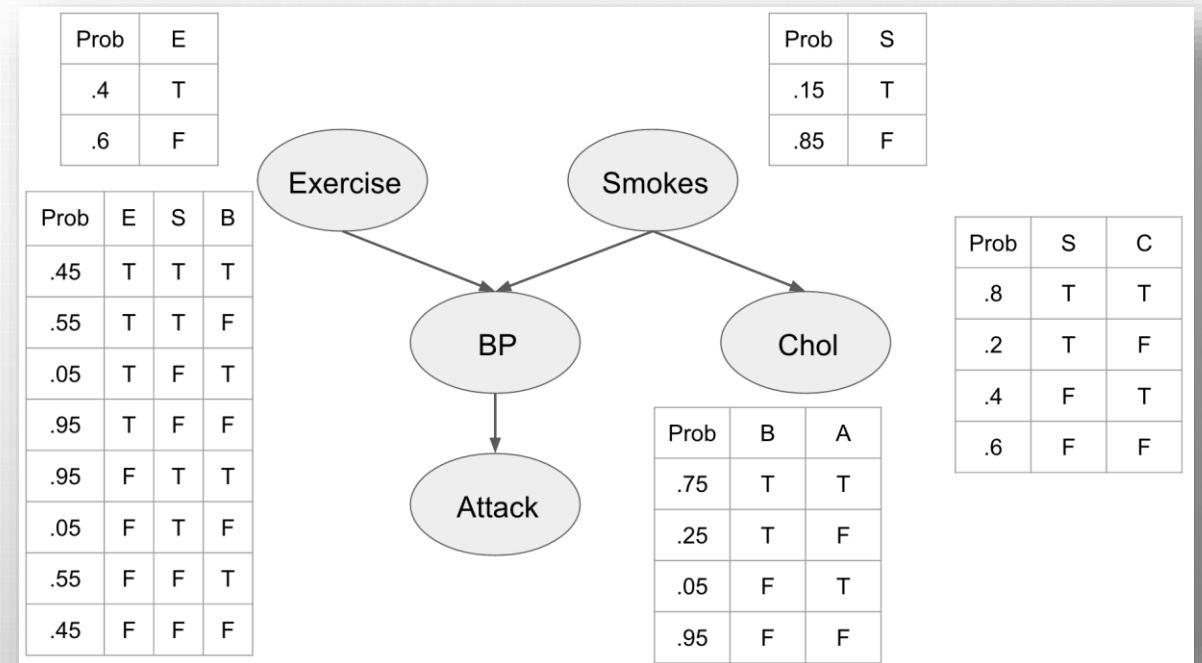


REDES BAYESIANAS

Representam relações de **dependência probabilística** entre **variáveis** utilizando **grafos direcionados acíclicos (DAGs)**.

Exemplo

Diagnóstico médico, onde os sintomas estão condicionados a doenças.



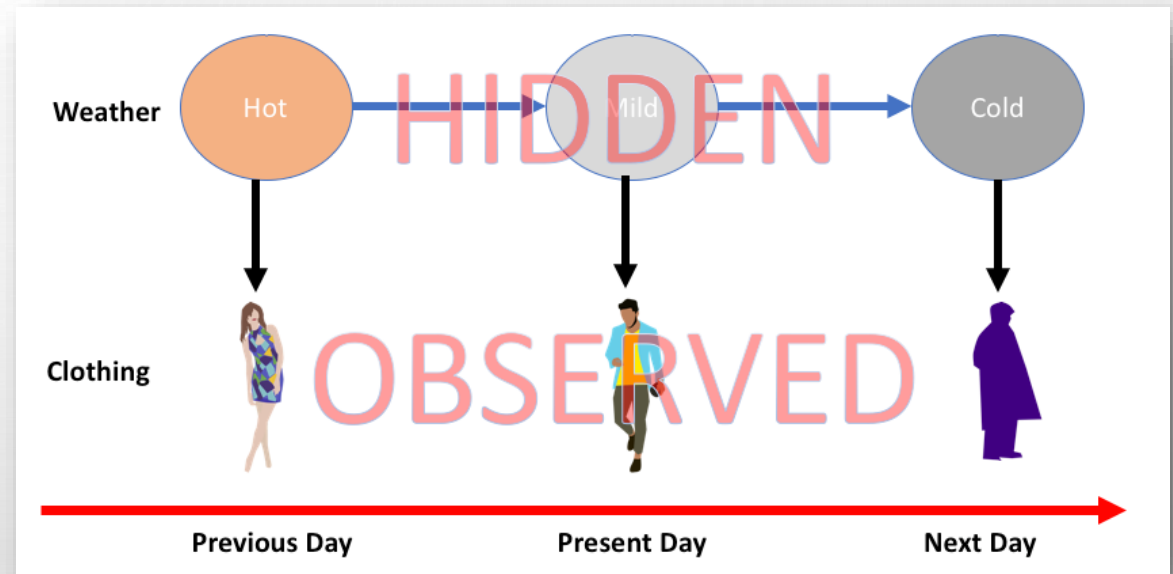
MODELO DE MARKOV OCULTO

Capturam relações temporais probabilísticas entre estados ocultos e observações. O modelo parte da premissa de que cada estado depende apenas do estado anterior, tornando o modelo eficiente para dados sequenciais.

Exemplos

Reconhecimento de fala

Processamento de linguagem natural

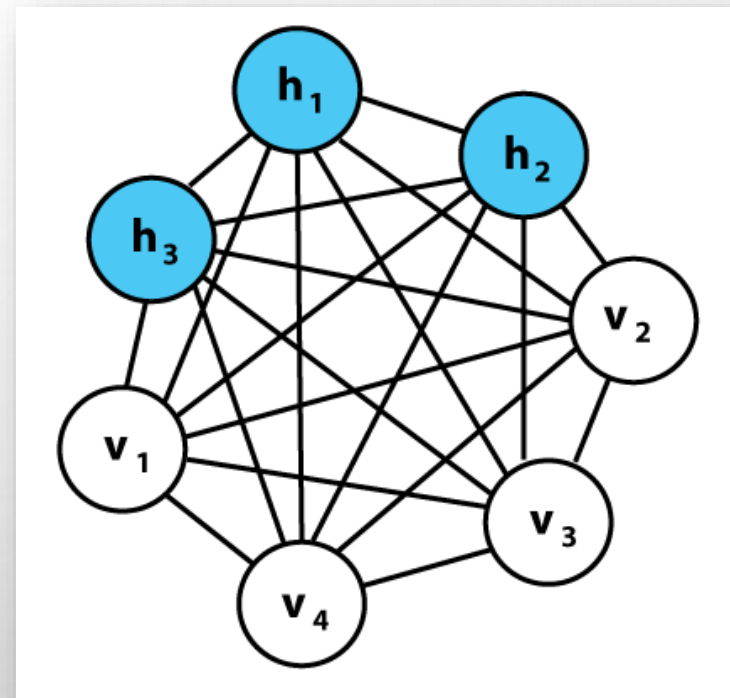


MÁQUINA DE BOLTZMANN 80'S

Redes neurais que tentam **inferir a distribuição probabilística** entre um **conjunto de variáveis**.

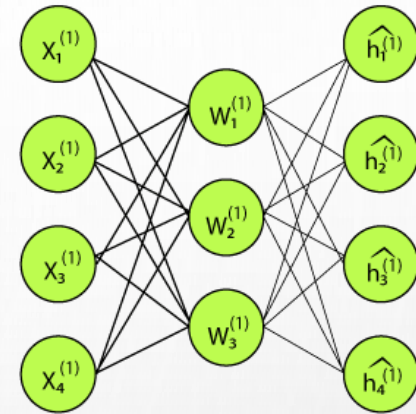
Funcionamento

- As Máquinas de Boltzmann (BM) são **totalmente conectadas e simétricas**, com **cada neurônio conectado a todos os outros**.
- São **treinadas para minimizar a energia** de uma configuração de estados (inspiradas na mecânica estatística).
- **Limitações:** Por ser totalmente conectada, o **treinamento de uma BM é computacionalmente caro e impraticável para redes grandes**.

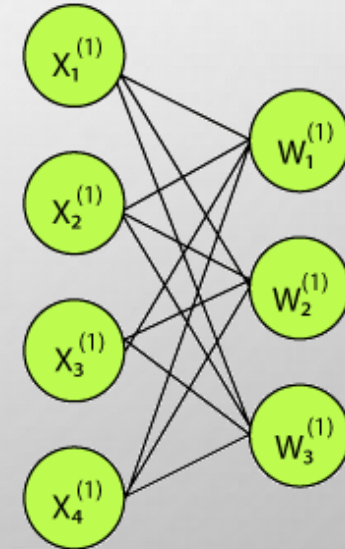


MÁQUINA DE BOLTZMANN RESTRITA 2000'S

RBM's são uma **simplificação das Máquinas de Boltzmann** com **restrições nas conexões** (nenhuma conexão entre neurônios da mesma camada).



AUTOENCODERS



RBM's

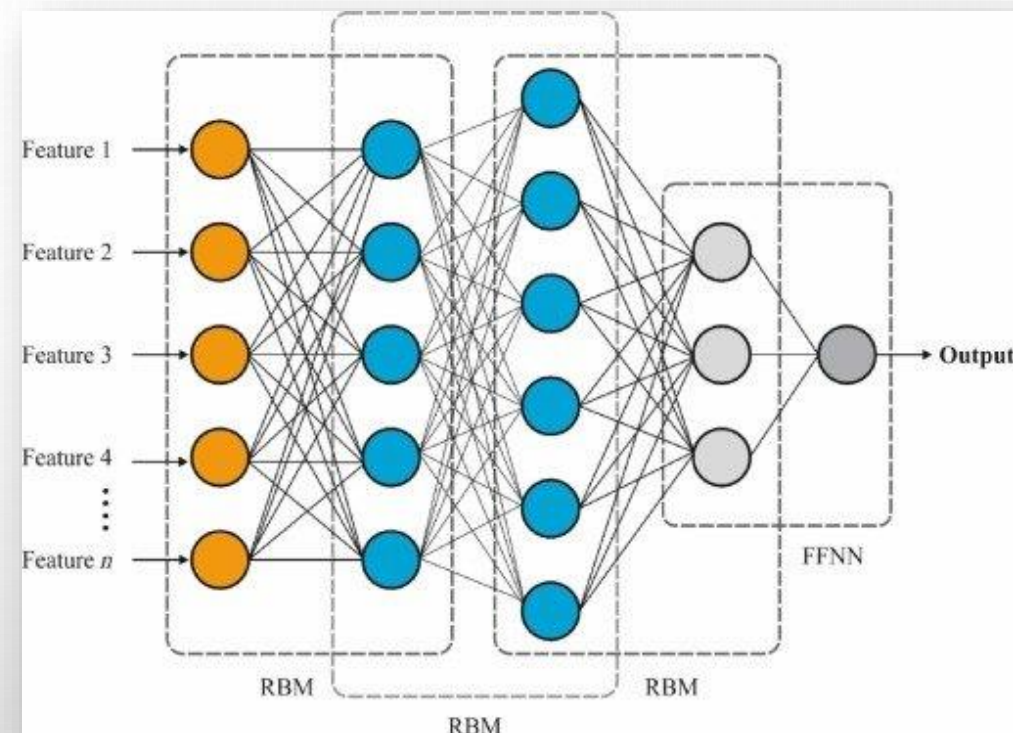
Características

- Camada visível (dados de entrada).
- Camada oculta (características latentes).
- Aprendizado: Algoritmo de **Contrastive Divergence** (Hinton, 2002) tornou o treinamento das RBMs viável em grandes datasets.

DEEP BELIEVE NETWORKS 2000'S

Deep Belief Networks (DBNs) são **redes neurais profundas** formadas por **múltiplas camadas empilhadas de RBMs**.

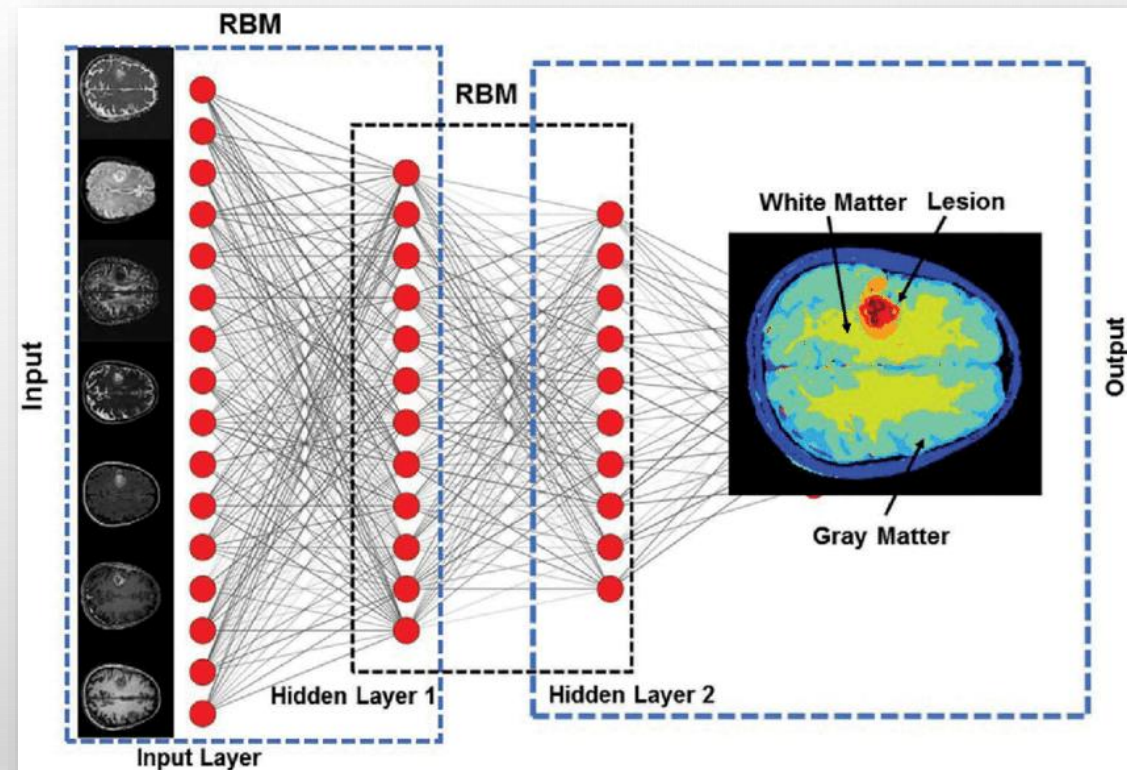
Cada camada é treinada individualmente de forma não supervisionada usando RBMs, e depois a rede é ajustada de maneira supervisionada para uma tarefa específica.



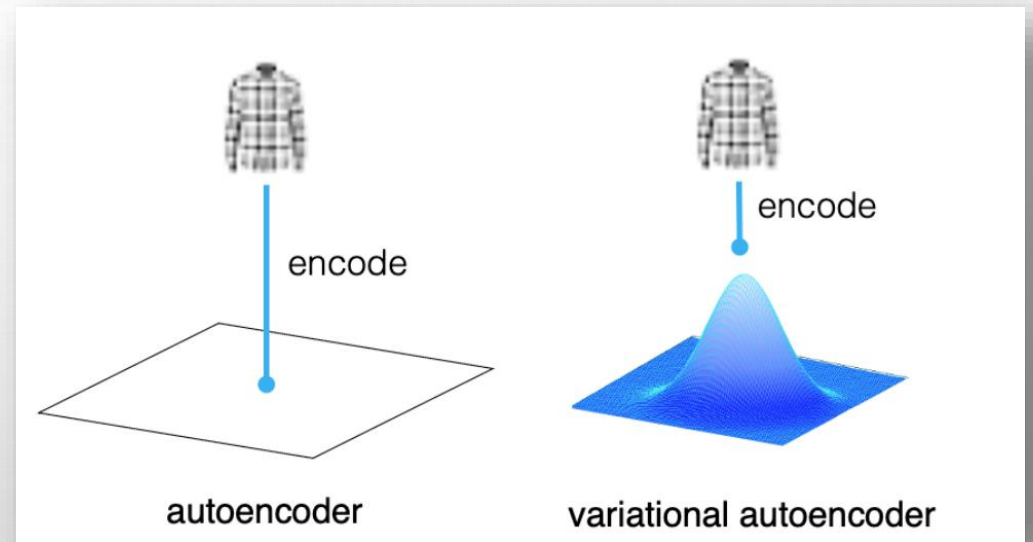
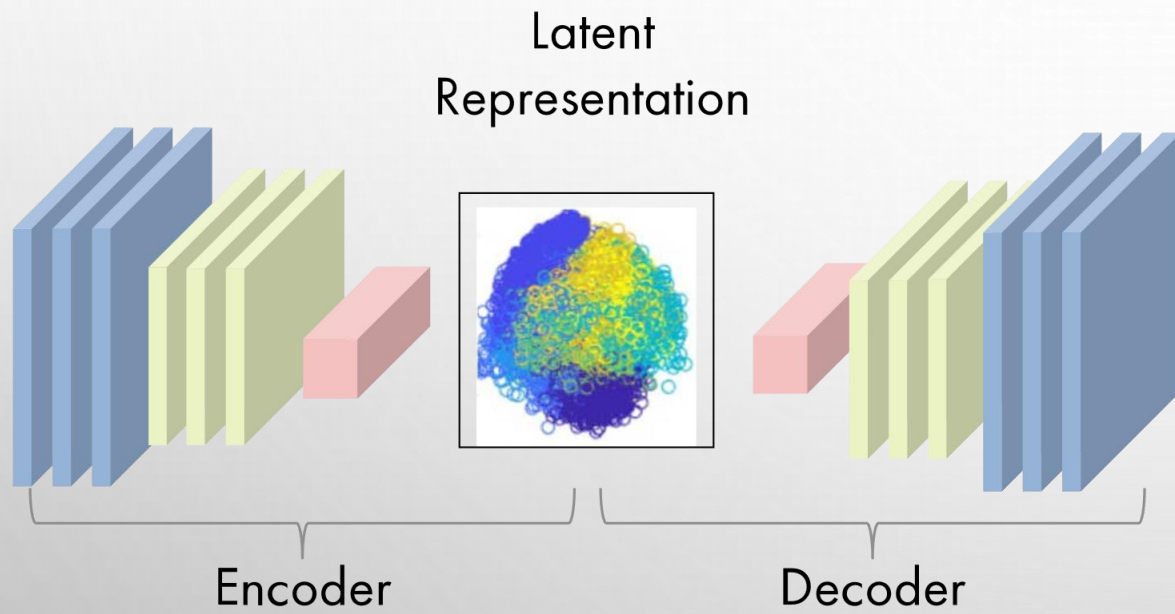
DEEP BELIEVE NETWORKS 2000'S

DBNs foram os primeiros modelos a **mostrar como redes neurais profundas poderiam ser treinadas com eficiência, solucionando o problema de dissipação do gradiente** que antes limitava o uso de redes profundas.

O sucesso dos DBNs trouxe de volta o interesse no uso de redes profundas, sendo precursor para arquiteturas mais avançadas.



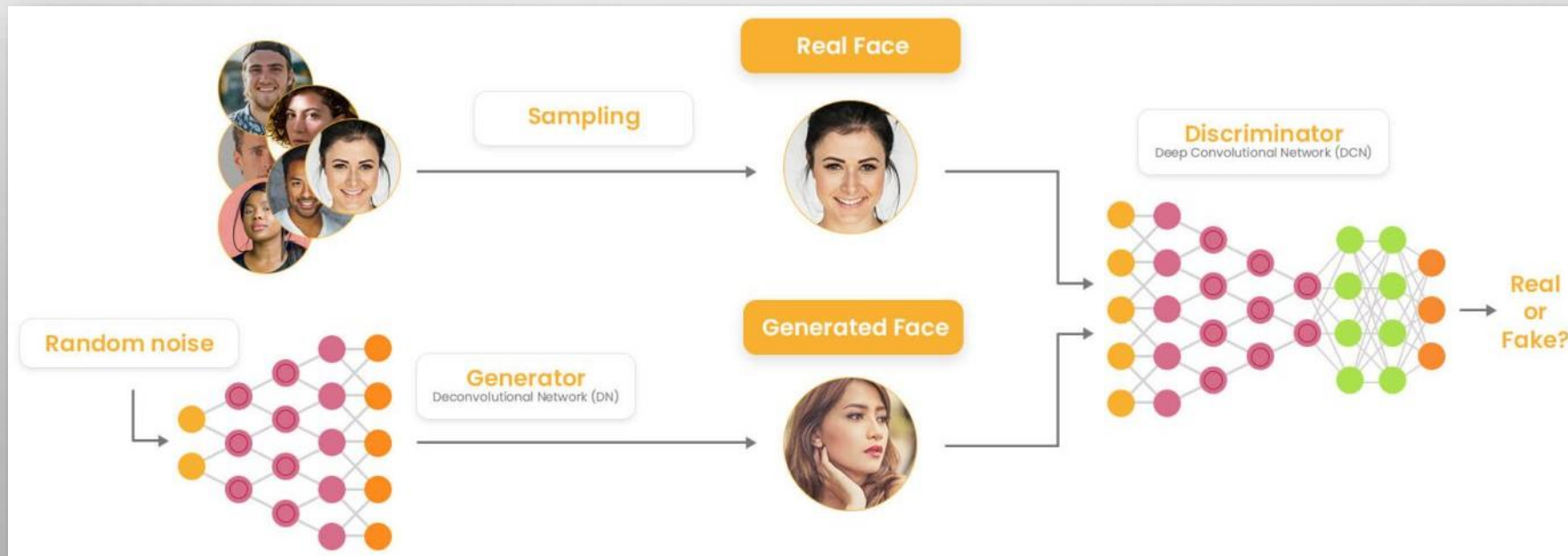
AUTOENCODERS



GENERATIVE ADVERSARIAL NETWORKS

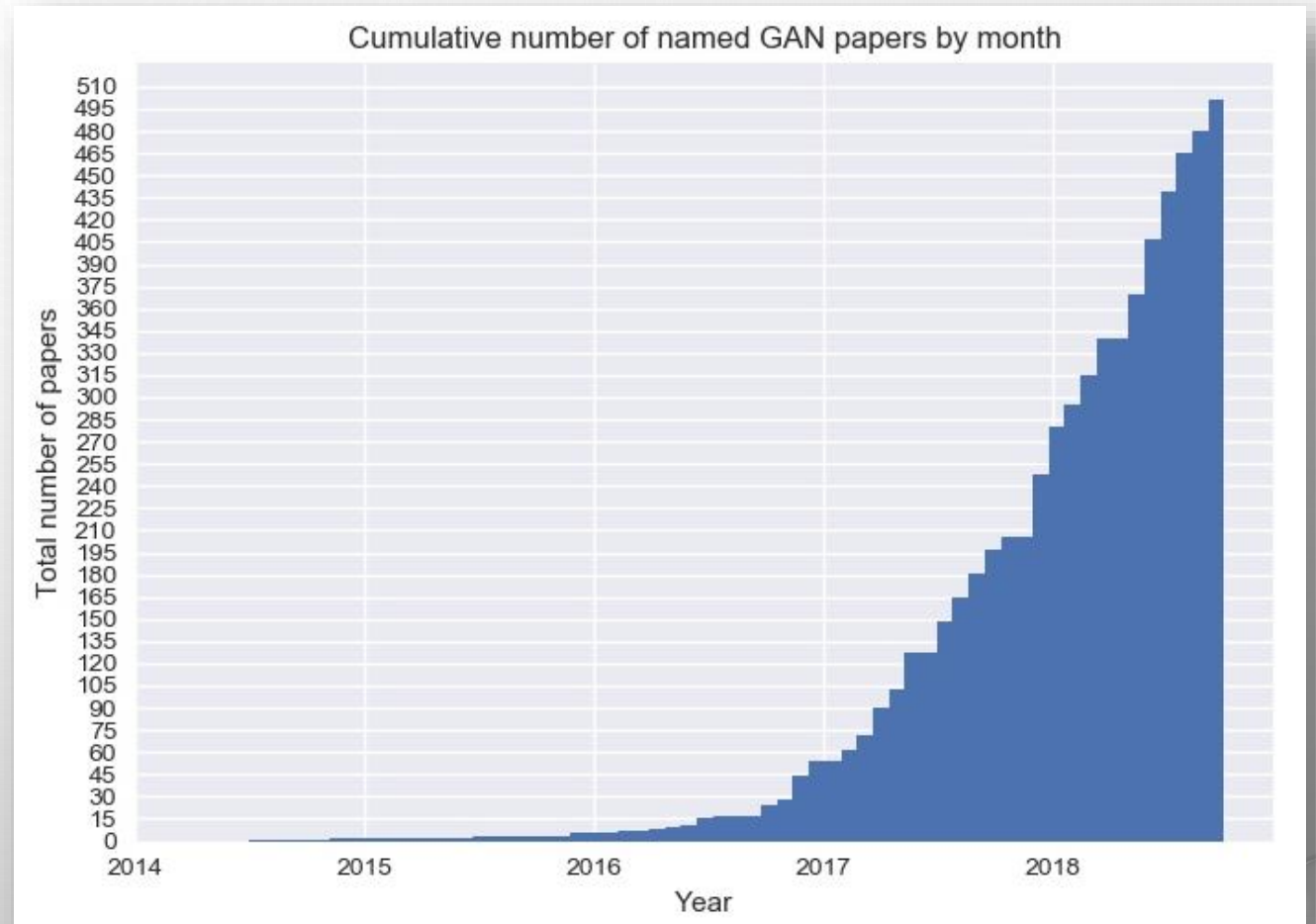
Generative Adversarial Networks (GANs) são redes neurais compostas por dois modelos: **um Gerador e um Discriminador**, que competem entre si em um **processo adversarial**.

Introduzidas por **Ian Goodfellow** em **2014**, GANs revolucionaram a geração de dados sintéticos.



TIPOS DE GAN

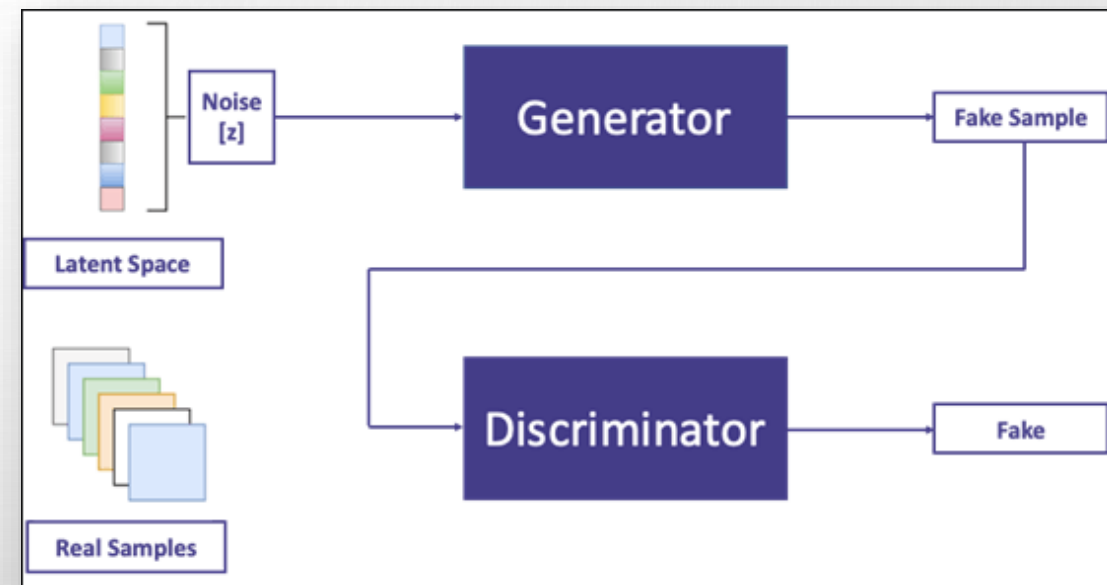
1. Vanilla GAN (2014)
2. Conditional GAN (cGAN) (2014)
3. Deep Convolutional GAN (DCGAN) (2015)
4. Wasserstein GAN (WGAN) (2017)
5. WGAN-GP (Wasserstein GAN with Gradient Penalty) (2017)
6. Least Squares GAN (LSGAN) (2017)
7. CycleGAN (2017)
8. Progressive GAN (PGGAN) (2017)
9. StyleGAN (2018)



VANILLA GAN (2014)

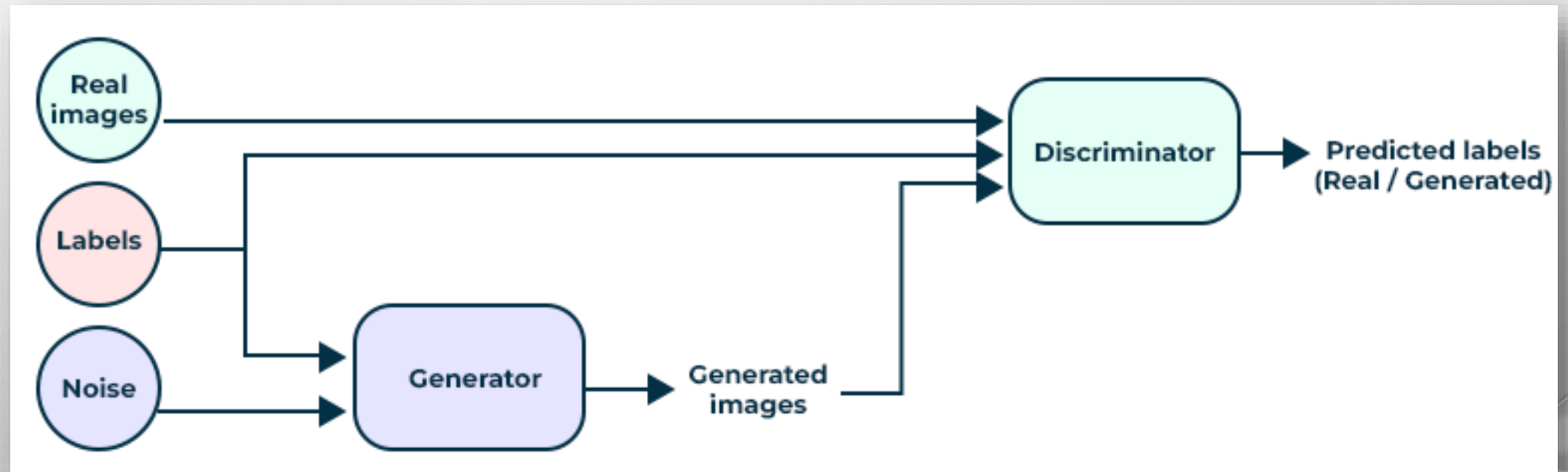
A arquitetura original de GANs, onde o **gerador** e o **discriminador** são simples redes feedforward.

Desvantagem: Dificuldade de treinamento devido à instabilidade e Mode Collapse.



CONDITIONAL GAN (2014)

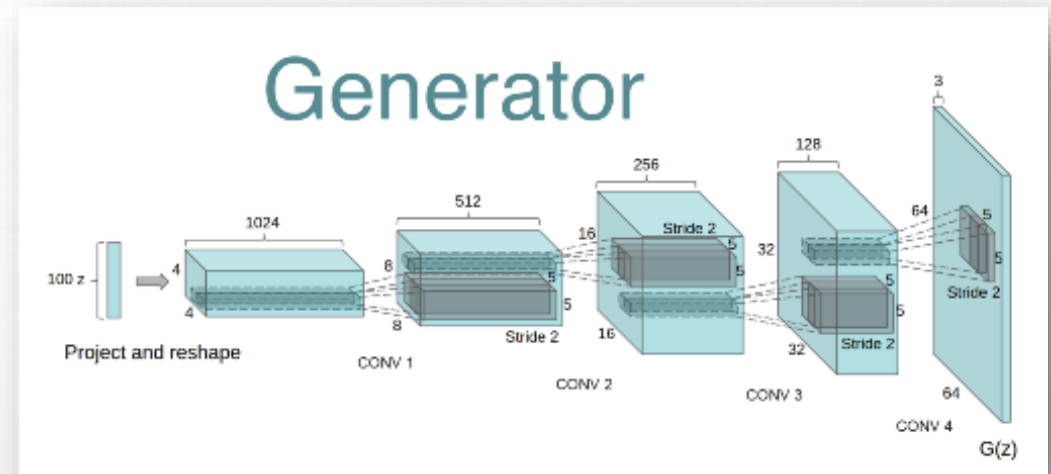
Adiciona uma **condição extra ao gerador e ao discriminador**, permitindo o **controle do tipo de amostra gerada**.



DEEP CONVOLUTIONAL GAN (2015)

Extensão das **GANs** com **camadas convolucionais profundas**, trazendo avanços significativos na **qualidade da geração de imagens**.

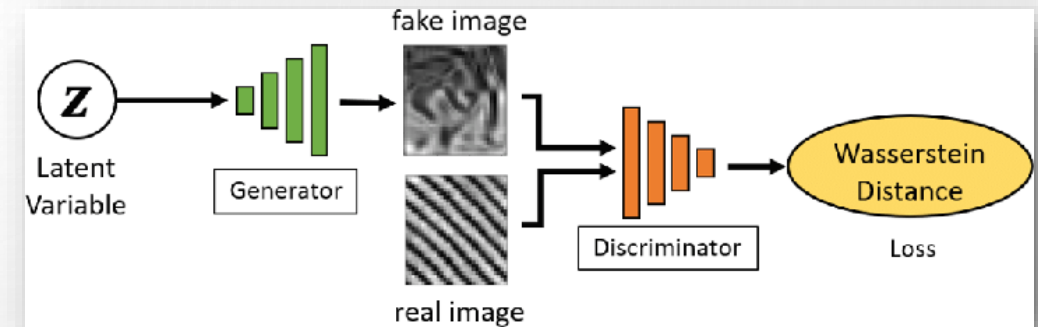
Usada amplamente em geração de imagens realistas e arte digital.



WASSERSTEIN GAN (2017)

Wasserstein GAN (WGAN)

- Resolve o problema de instabilidade no treinamento usando a distância de Wasserstein como função de perda.
- Vantagem: Melhora a estabilidade e evita mode collapse, tornando o treinamento mais confiável.



WGAN-GP (Wasserstein GAN with Gradient Penalty)

- Aperfeiçoamento do WGAN que adiciona uma penalidade de gradiente, melhorando ainda mais a estabilidade do treinamento.

$$W_p(\mu, \nu) := \left(\inf_{\gamma \in \Gamma(\mu, \nu)} \int_{M \times M} d(x, y)^p d\gamma(x, y) \right)^{1/p}$$

LEAST SQUARES GAN (2017)

Modifica a função de perda para minimizar a divergência entre as distribuições geradas e reais, resultando em imagens de maior qualidade.



(a) LSGANs.



(b) Regular GANs.



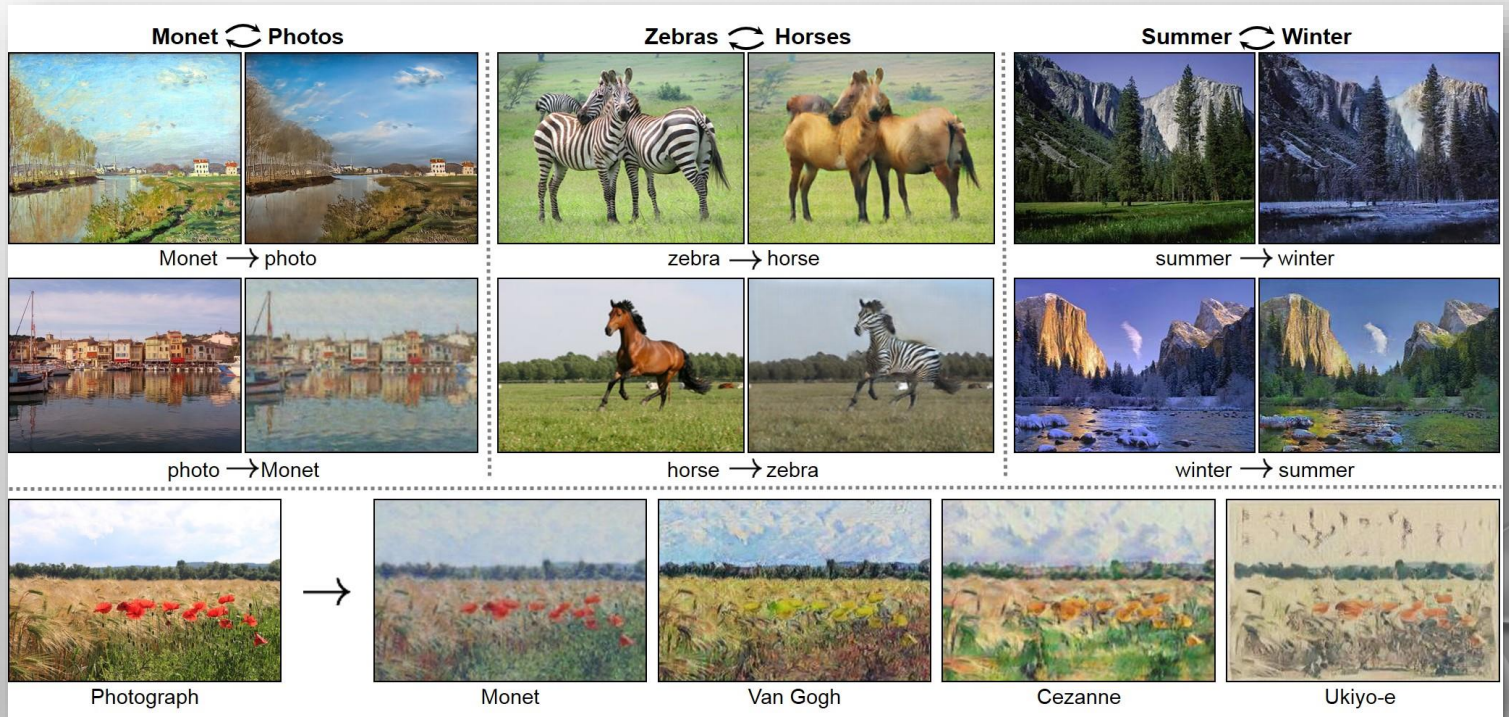
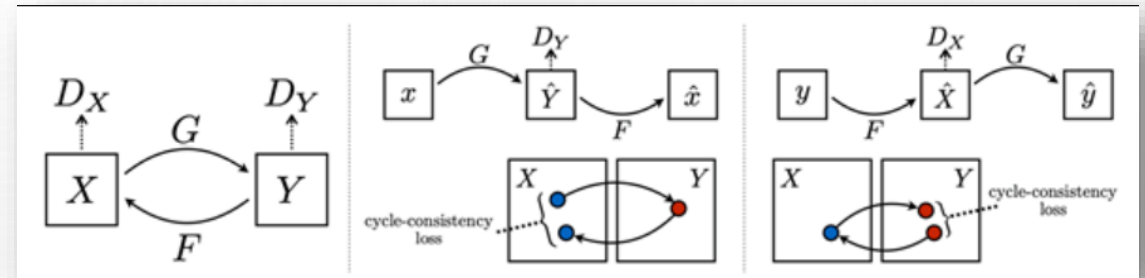
(c) LSGANs.



(d) Regular GANs.

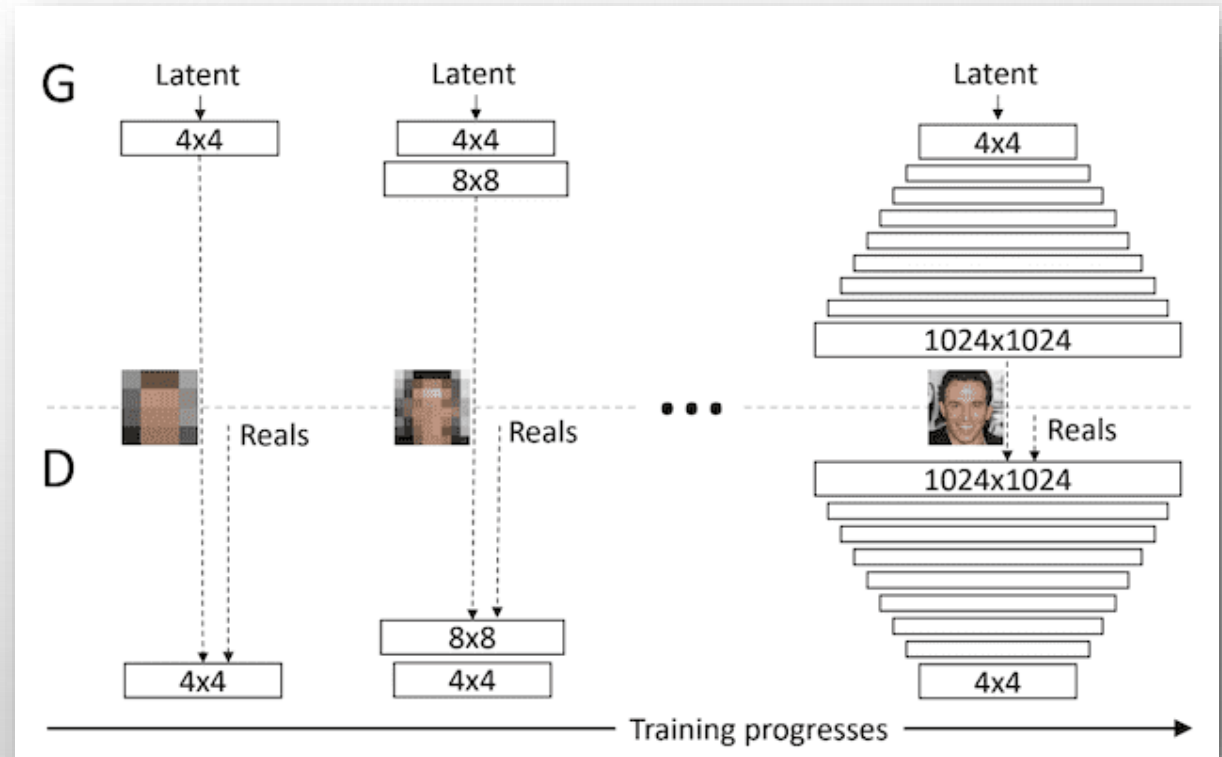
CYCLE GAN (2017)

Permite a **transformação**
de **imagens** **entre**
domínios **sem** **a**
necessidade **de** **pares** **de**
imagens **correspondentes.**



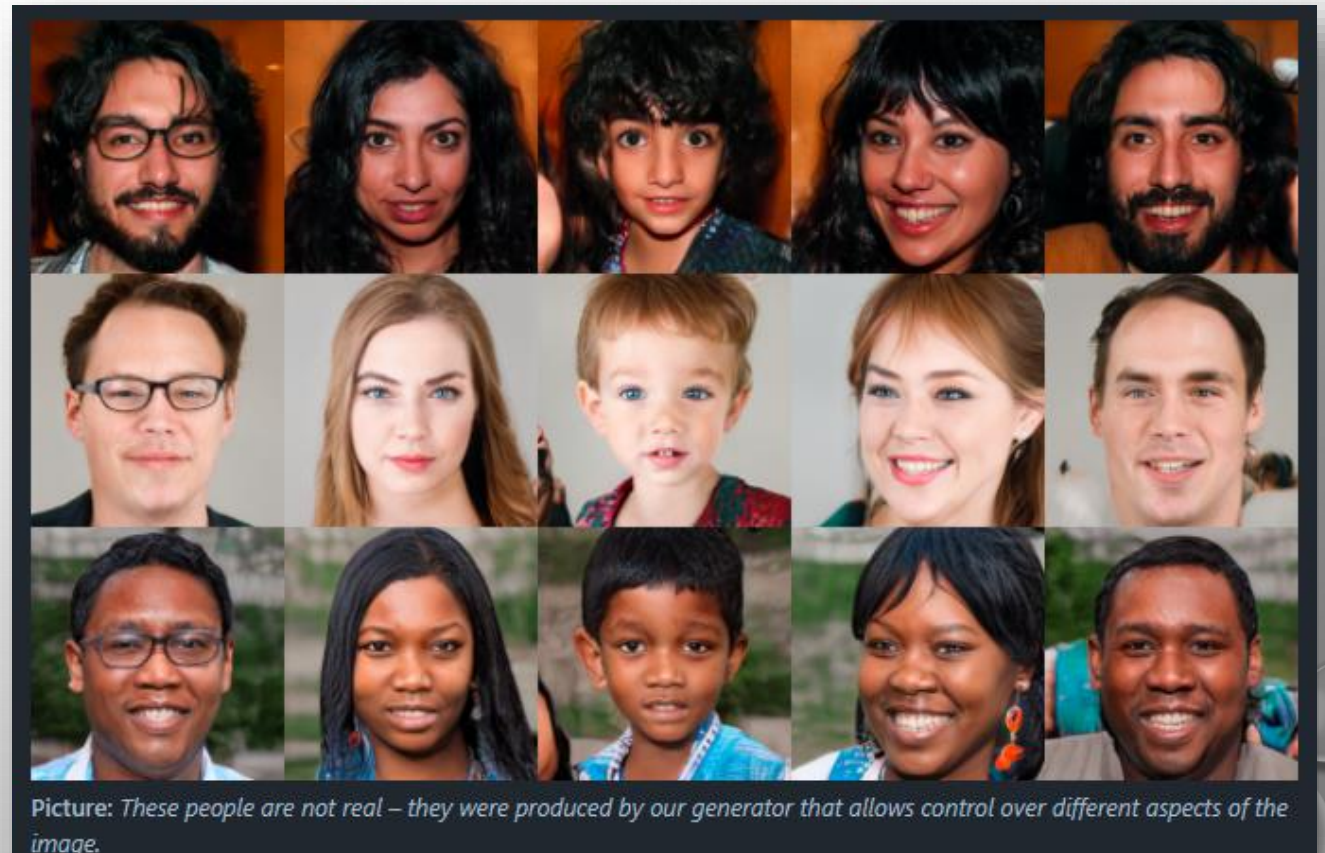
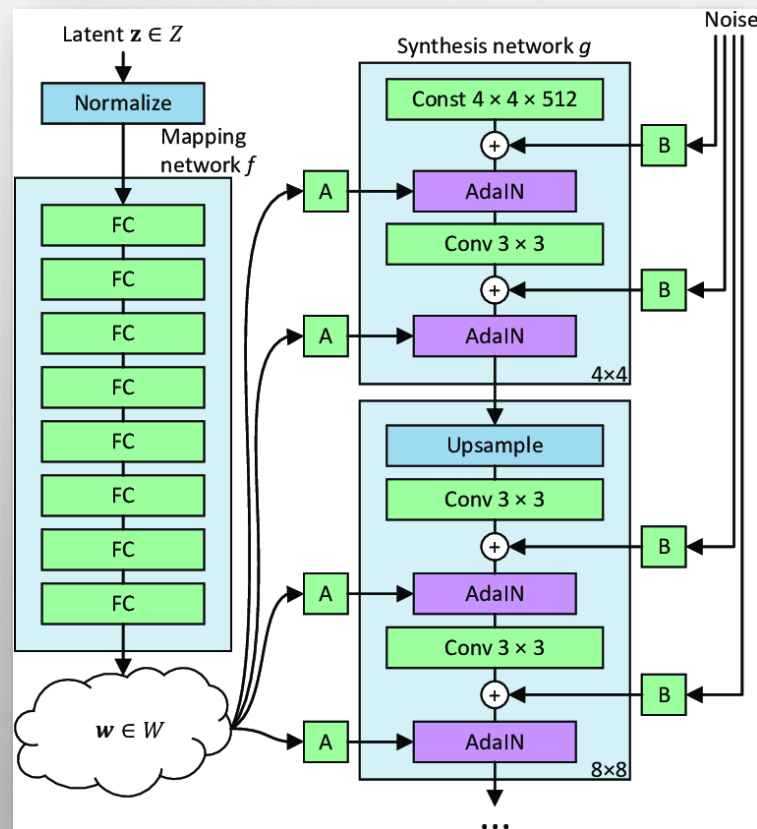
PROGRESSIVE GAN (2017)

Treinamento progressivo, começando com **imagens de baixa resolução** e **aumentando gradualmente a resolução** ao longo do tempo, resultando em imagens mais estáveis e de alta qualidade.



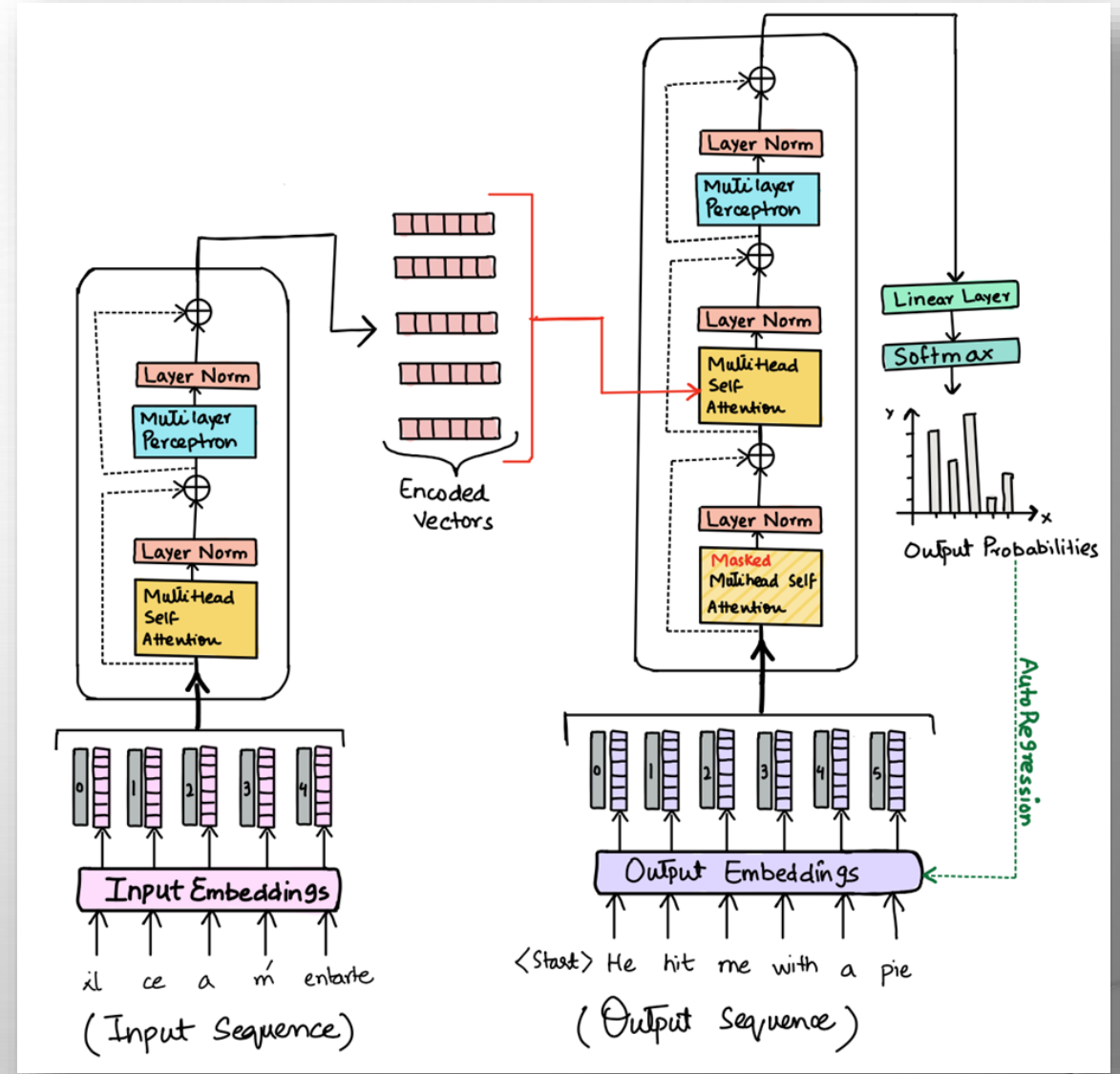
STYLE GAN (2018)

Desenvolvida pela NVIDIA, é uma das arquiteturas mais avançadas para geração de imagens realistas, com controle detalhado sobre os atributos estilísticos das imagens geradas.



TRANSFORMERS (2017)

A arquitetura de Transformers, introduzida no artigo "Attention is All You Need", **revolucionou o processamento de linguagem natural (NLP)** ao **abandonar redes recorrentes e convolucionais** em favor do **mecanismo de atenção**.

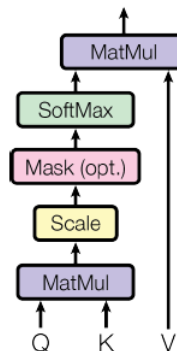


TRANSFORMERS (2017)

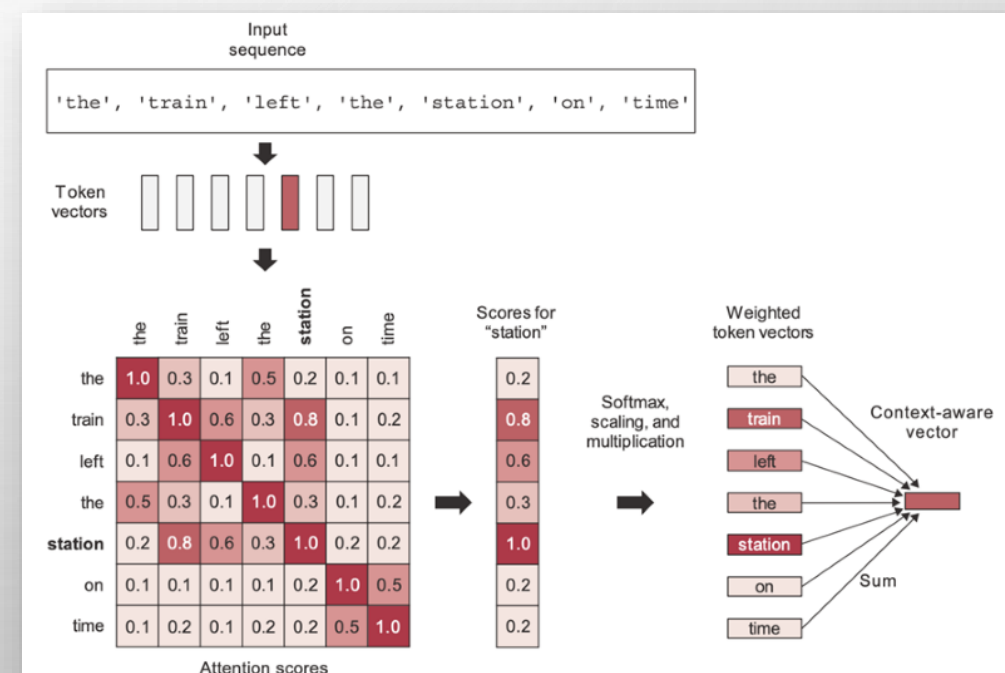
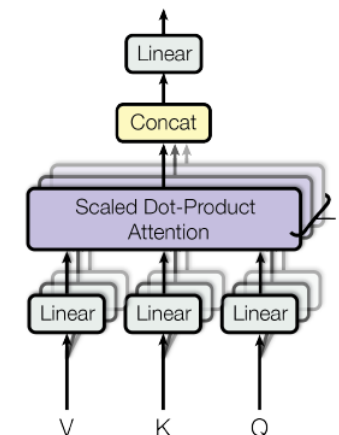
Como Funciona o Self-Attention?

Imagine que estamos processando uma frase, e o modelo precisa prever a próxima palavra ou entender o significado de uma palavra dentro da frase. O **self-attention** permite que cada palavra se relacione com todas as outras palavras da sequência de entrada, o que ajuda o modelo a **determinar quais palavras são mais relevantes para entender ou prever a palavra atual**.

Scaled Dot-Product Attention

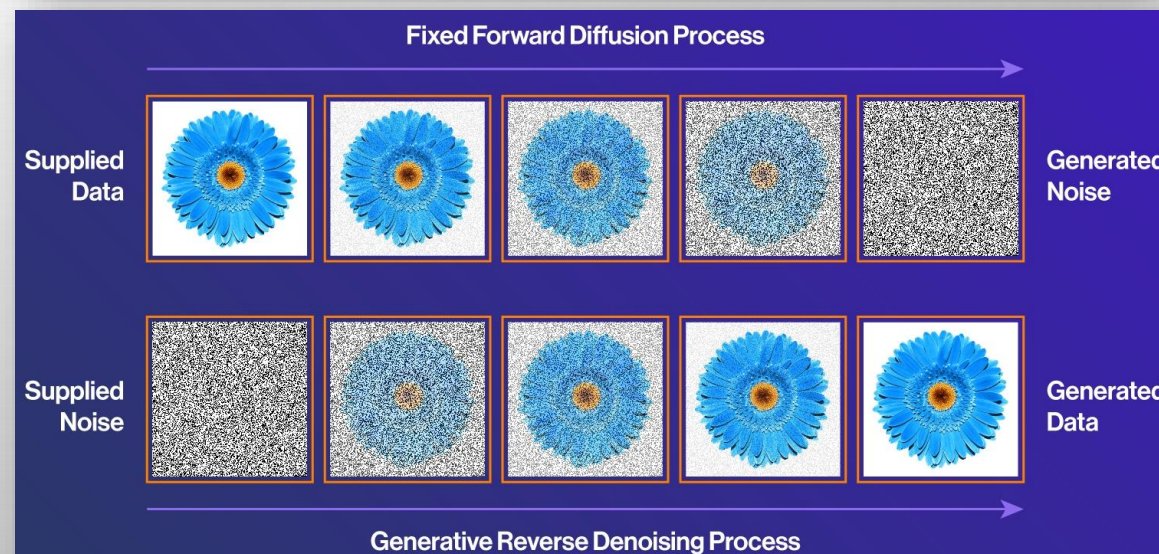
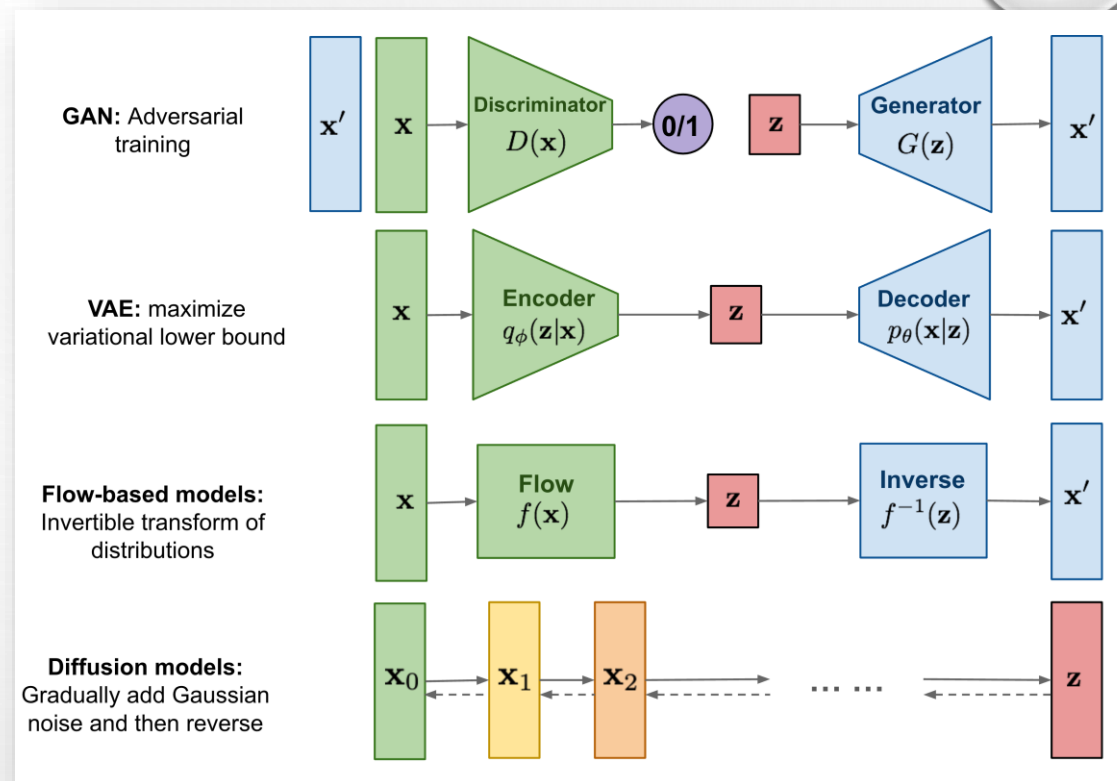


Multi-Head Attention



DIFUSÃO (2021)

Modelos de Difusão são uma nova classe de modelos generativos baseados em um **processo de difusão**. O conceito central é que os dados são **gradualmente corrompidos** com **ruído gaussiano**, e o modelo é treinado para reverter esse processo, reconstruindo os dados originais a partir de versões ruidosas.



WORKSHOP

- ~~• QUAL A TOPOLOGIA DE DEEP LEARNING ADEQUADA PARA O MEU TRABALHO?~~
- ~~• QUAL CAPÍTULO DO LIVRO MELHOR SE ENQUADRA NO MEU TRABALHO?~~
- ~~• AULA 3: NOVO CICLO DE BUSINESS UNDERSTANDING / GRUPO + MODELO BASELINE TREINADO~~
- **AULA 5 OU 7: MODELO PROFUNDO TREINADO**
- **AULA 7: DEPLOYMENT DO MODELO***
- **AULA 3-7 > APRESENTAÇÃO TEÓRICA DA(S) TOPOLOGIA(S) + LEITURA DE ARTIGO + ACOMPANHAMENTO DOS TRABALHOS + DEEP DIVE NO CÓDIGO (POR GRUPO)**
- **APRESENTAÇÃO FINAL DOS TRABALHOS**

**Passo Opcional*